Please substitute the following paragraphs of the **SPECIFICATION** for those currently of record:

[33] Digital Karyotyping was performed on DNA from colorectal cancer cell lines DiFi and Hx48, and from lymphoblastoid cells of a normal individual (GM12911, obtained from Coriell Cell Repositories, NJ). Genomic DNA was isolated using DNeasy or QIAamp DNA blood kits (Qiagen, Chatsworth, CA) using the manufacturers' protocols. For each sample 1 μg of genomic DNA was sequentially digested with mapping enzyme *Sac*I, ligated to 20-40 ng of biotinylated linker (5'-biotin-TTTGCAGAGGTTCGTAATCGAGTTGGGTGAGCT(SEQ ID NO: 1), 5'-phosphate-CACCCAACTCGATTACGAACCTCTGC-3' (SEQ ID NO: 2)) (Integrated DNA Technologies, Coralville, IA) using T4 ligase (Invitrogen, Carlsbad, CA), and then digested with the fragmenting enzyme *Nla*III. DNA fragments containing biotinylated linkers were isolated by binding to streptavidin coated magnetic beads (Dynal, Oslo, Norway). The remaining steps were similar to those described for LongSAGE of cDNA (18). In brief, linkers containing *Mme*I recognition sites were ligated to captured DNA fragments, and tags were released with *Mme*I (University of Gdansk Center for Technology Transfer, Gdansk, Poland and New England Biolabs, Beverly, MA). The tags were ligated to form ditags, and the ditags were isolated and then ligated to form concatemers which were cloned into pZero (Invitrogen, Carlsbad, CA). Sequencing of concatamer clones was performed using the Big Dye Terminator v3.0 Kit (Applied Biosystems) and analyzed with a SCE-9610 192- capillary electrophoresis system (SpectruMedix, State College, PA) or by contract sequencing at Agencourt (Beverly, MA). Digital Karyotyping sequence files were trimmed using Phred sequence analysis software (CodonCode, MA) and 21 bp genomic tags were extracted using the SAGE2000 software package, which identifies the fragmenting enzyme site between ditags, extracts intervening tags, and records them in a database. Detailed protocols for performing Digital Karyotyping and software for extraction and analysis of genomic tags are available at ~~http://www.sagenet.org/digital_karyotyping.htm~~ the http protocol, the sagenet.org domain name, and the digital_karyotyping.htm path.

[35] All tags adjacent to the *Nla*III fragmenting enzyme (CATG) sites closest to *Sac*I mapping enzyme sites were computationally extracted from the human genome sequence (UCSC June 28, 2002 Assembly, ~~http://genome.ucsc.edu/~~ the http protocol, the genome.ucsc.edu domain name). Of the 1,094,480 extracted tags, 730,862 were obtained from unique loci in the genome and were termed virtual tags. The experimentally derived genomic tags obtained from NLB, DiFi and Hx48 cells were electronically matched to these virtual tags. The experimental tags with the same sequence as virtual tags were termed filtered tags and were used for subsequent analysis. The remaining tags corresponded to repeated regions, sequences not present in the current genome database release, polymorphisms at the tag site, or sequencing errors in the tags or in the genome sequence database. Tag densities for sliding windows containing N virtual tags were determined as the sum of experimental tags divided by the average number of experimental tags in similar sized windows throughout the genome. Tag densities were dynamically analyzed in windows ranging from 50 to 1000 virtual tags. For windows of 1000 virtual tags, DiFi tag densities were normalized to evaluated NLB tag densities in the same sliding windows to account for incomplete filtering of tags matching repetitive sequences, and visualized using tag density maps. For windows less than 1000 virtual tags, a bitmap viewer was developed that specifically identified tag densities above or below defined thresholds.

[36] Genome content differences between DiFi and normal cells were determined by quantitative real-time PCR using an iCycler apparatus (Bio-Rad, Hercules, CA). DNA content was normalized to that of Line-1, a repetitive element for which copy numbers per haploid genome are similar among all human cells (normal or neoplastic). Copy number changes per haploid genome were calculated using the formula $2^{(Nt-Nline)-(Dt-Dline)}$ where N_t is the threshold cycle number observed for experimental primer in the normal DNA sample, N_{line} is the threshold cycle number observed for Line-1 primer in the normal DNA sample, D_t is the average threshold cycle number observed for the experimental primer in DiFi and D_{line} is the average threshold cycle number observed for Line-1 primer in DiFi. Conditions for amplification were: one cycle of 94° C, 2 min followed by 50 cycles of 94° C, 20 sec, 57° C, 20 sec, 70° C, 20 sec. Threshold cycle numbers were obtained using iCycler software v2.3. PCR reactions for each primer set were

performed in triplicate and threshold cycle numbers were averaged. For analysis of homozygous deletions, presence or absence of PCR products was evaluated by gel electrophoresis. PCR primers were designed using Primer 3 (~~http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi~~ the http protocol, the www-genome.wi.mit.edu domain name, the path cgi-bin/primer/primer3_www.cgi) to span a 100 – 200 bp non-repetitive region and were synthesized by GeneLink (Hawthorne, NY). Primer sequences for each region analyzed in this study are included in Table 4.